CSIRO PUBLISHING

Invertebrate Systematics, 2011, **25**, 322–333
http://dx.doi.org/10.1071/IS11022

Barcoding type specimens helps to identify synonyms and an unnamed new species in *Eumunida* Smith, 1883 (Decapoda : Eumunididae)

Nicolas Puillandre[A,E], *Enrique Macpherson*[B], *Josie Lambourdière*[C], *Corinne Cruaud*[D], *Marie-Catherine Boisselier-Dubayle*[A,C] *and Sarah Samadi*[A,C]

[A]'Systématique, Adaptation et Evolution', UMR 7138 UPMC-IRD-MNHN-CNRS (UR IRD 148), Muséum national d'Histoire naturelle, Département Systématique et Evolution, CP 26, 57 Rue Cuvier, F-75231 Paris Cedex 05, France.
[B]Centro de Estudios Avanzados de Blanes (CEAB-CSIC), C. acc. Cala Sant Francesc 14, 17300 Blanes, Girona, Spain.
[C]Service de systématique moléculaire (CNRS-MNHN, UMS2700), Muséum national d'Histoire naturelle, Département Systématique et Evolution, CP 26, 57 Rue Cuvier, F-75231 Paris Cedex 05, France.
[D]GENOSCOPE, Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry Cedex France.
[E]Corresponding author. Email: puillandre@mnhn.fr

Abstract. The primary purpose of DNA-barcoding projects is to generate an efficient expertise and identification tool. This is an important challenge to the taxonomy of the 21st century, as the demand increases and the expert capacity does not. However, identifying specimens using DNA-barcodes requires a preliminary analysis to relate molecular clusters to available scientific names. Through a case study of the genus *Eumunida* (Decapoda : Eumunididae), we illustrate how naming molecule-based units, and thus providing an accurate DNA-based identification tool, is facilitated by sequencing type specimens. Using both morphological and unlinked molecular markers (COI and 28S genes), we analysed 230 specimens from 12 geographic areas, covering two-thirds of the known diversity of the genus, including type specimens of 13 species. Most hypotheses of species delimitation are validated, as they correspond to molecular units linked to only one taxonomic name (and *vice versa*). However, a putative cryptic species is also revealed and three entities previously named as distinct species may in fact belong to a single one, and thus need to be synonymised. Our analyses, which integrate the current naming rules, enhance the α-taxonomy of the genus and provide an effective identification tool based on DNA-barcodes. They illustrate the ability of DNA-barcodes, especially when type specimens are included, to pinpoint where a taxonomic revision is needed.

Received 25 May 2011, accepted 9 November 2011, published online 22 December 2011

Introduction

When describing a new species, the taxonomists provide a species name and designate one (or several) type specimens to which this name is permanently attached. A species name allows us to designate a testable species hypothesis, and the type specimens provide the link with the name of this hypothesis. Designating species hypotheses by species names allows anyone to associate newly examined specimens to already proposed species hypotheses. However, proposing species hypotheses, species names and species identifications are three distinct tasks that should not be confused (Dayrat 2006). They can be distinguished as follows: (1) the scientific task consists of proposing hypotheses about species boundaries, based on the comparison of characters or on biological criteria; (2) the naming task deals with assigning names to such species hypotheses; and (3) the identification task is to identify specimens in the light of already named species hypotheses.

Within this methodological framework, the primary purpose of DNA barcoding projects is not to produce new taxonomic hypotheses and to name them – Tasks 1 and 2 – but to facilitate taxonomic identification – Task 3 – by developing a global standard for the identification of biological species based on molecular data (Hebert and Gregory 2005; Schindel and Miller 2005). However, identifying specimens using only their barcode sequences requires a database that includes the sequences and the corresponding specimen data, authoritatively identified using morphological characters. Furthermore, a prior analysis of the molecular diversity of the groups is necessary to confirm (or reject) that DNA barcodes may be used as a diagnostic character for the species at hand, i.e. that intraspecific and interspecific genetic distances are separated by a 'barcode gap'. In that way, the identification of new specimens using such a DNA library would follow the opinion of the taxonomist that has identified the specimens of the DNA barcode library. Here, two problems

need to be addressed. First, a link between DNA-based species hypotheses and already available morphological species hypotheses (and thus species names) needs to be assessed. For example, in the study of Smith *et al*. (2007), it was not possible to ascertain the link between genetic clusters and available names with full confidence because no DNA barcode was obtained for the holotype; this uncertainty in the assignation of species names to species hypotheses was shown by indicating the scientific names in quotation marks. Second, one important by-product of DNA barcoding as an identification tool for taxonomy is the detection of specimens that cannot be attributed to any available species hypothesis, and for which a new hypothesis – and thus a new name – may be proposed (e.g. Padial and De La Riva 2007). Once again, the attribution of available species names to genetic clusters is critical to clearly highlight genetic clusters that would deserve a new species name. Thus, because DNA barcodes can be used both to attribute species names to a given specimen and to flag genetic clusters for which no name is available, we should clarify how names are – or should be – given to species hypotheses. This can be achieved by the sequencing of type specimens.

Using a case study of the genus *Eumunida* Smith, 1883 (Decapoda : Chirostyloidea : Eumunididae), we illustrate here the difficulties of this naming task, in the context of the development of DNA barcodes as an identification tool. We selected this genus because most species have been described recently and the conservation of name-bearing specimens in the collections allows us to access molecular characters. Many species were described using material that has been preserved in 70% ethanol, the samples are housed in the collection of the Museum national d'Histoire naturelle in Paris, having been collected over a quarter of a century's exploration in the south-west Pacific (Bouchet *et al*. 2008). In this case study we integrate the three tasks of taxonomy. Our specific aims are thus: (1) to test the robustness of recognised species hypotheses and, if needed, to propose new ones; and (2) to name the revised set of species hypotheses. This way, the efficiency of DNA barcodes as an identification key will also be evaluated. To that end, we gathered mitochondrial and nuclear data for 230 specimens attributed to the genus *Eumunida*, including type specimens, for a large proportion of the described species. We also compared the distribution of morphological characters used in the identification keys over the identified genetic clusters. The inclusion of type specimens in the dataset unambiguously links genetic clusters to taxon names.

Materials and methods

Material and DNA sequencing

From the collections of the Museum national d'Histoire naturelle, Paris (MNHN) we selected 230 specimens of *Eumunida* from the South West Pacific and Indian Oceans (Table 1). Among them, nine are holotypes and 24 are paratypes, representing 13 different species. The 197 remaining specimens were morphologically identified to species level and attributed to 17 valid names of eumunid species. Thus, more than half of the species diversity currently recognised in the genus *Eumunida* is represented in our dataset (Tables 2, 3). These 17 species hypotheses are represented by 1–95 specimens, with an average of 12.05 specimens per species (Table 1). These morphological identifications were used as primary species hypotheses. The morphological characters used in species identification for all the species in the genus were listed and used to build a morphological matrix (Tables 2, 3).

DNA was extracted from a piece of muscle tissue using the DNeasy® 96 Tissue kit (Qiagen), and specimens were kept as vouchers. Fragments of the Cytochrome Oxydase I (COI) mitochondrial gene and 28S rDNA nuclear gene were amplified using universal primers LCO1490 (5′-GGTCAA CAAATCATAAAGATATTGG-3′) and HCO2198 (5′-TAA ACTTCAGGGTGACCAAAAAATCA-3′) (Folmer *et al*. 1994), and C1′ (5′- ACCCGCTGAATTTAAGCAT-3′ (Jovelin and Justine 2001) and D2 (5′-TCCGTGTTTCAAGACGG-3′ (Dayrat *et al*. 2001). All PCR reactions were performed in 25 µL, containing 3 ng of DNA, 1X reaction buffer, 2.5 mM $MgCl_2$, 0.26 mM dNTP, 0.3 µM of each primer, 5% DMSO and 1.5 units of Q-Bio Taq, QBiogene for COI gene and Taq Core Kit 2, QBiogene for 28S rDNA gene. Thermocycles consisted of an initial denaturation step at 94°C for 4 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 48°C for COI gene and 56°C for 28S rDNA gene for 40 s and extension at 72°C for 1 min. The final extension was at 72°C for 10 min. Some PCR products were purified using Montage™ PCR Centrifugal Filter Devices (Millipore) and sequenced on a CEqn 2000™ automated sequencer (Beckman) – corresponding to GenBank accession numbers AY800009–800046, AY800048, AY800050, AY800051, AY800055–800065 and DQ011181–011220. The other PCR products were purified and sequenced by the Genoscope (GenBank accession numbers EU243337–EU243562 for COI gene and EU243574–EU243663 for 28S rDNA gene). In all cases, both directions were sequenced to confirm accuracy of each haplotype sequences.

Phylogenetic analyses

Sequences were manually aligned for the COI gene, and the Clustal W algorithm (default parameters) implemented in BioEdit (Hall 1999) was used for alignment of our 28S rDNA sequences. Since all the species analysed here belong to a single genus, the sequence variability and the number of gaps for the 28S gene were reduced. Consequently, we considered that homology was confidently inferred using Bioedit. The RNAalifold webserver (http://rna.tbi.univie.ac.at/cgi-bin/RNAalifold.cgi) was used to predict a consensus secondary structure for the 28S gene and to identify the loops and stems. Loops generally correspond to variable regions, as opposed to stems, which are generally more conserved. In consequence, two different models of evolution were used for the phylogenetic analyses of the 28S data. Best-fit models of evolution were selected for the COI genes and for the loops and stems partitions of the 28S gene using Modellgenerator V.85 (Keane *et al*. 2006) under the Bayesian Information Criterion, with four discrete gamma categories. The best-fit models of evolution are the HKY+I+G (with $I = 0.6$ and $\alpha = 0.62$) for the COI gene, the TrNef+I+G ($I = 0.31$, $\alpha = 0.15$) for the 28S gene, the K80+G ($\alpha = 0.5$) for the loops of the 28S gene and the K80+G ($\alpha = 0.25$) for the stems of the 28S gene.

As distances-based methods are classically used in barcode studies, a genetic distance matrix including all sequences was

Table 1. Description of the specimens analysed in this study

MNHN ID	Geographic area	Morphological ID	Status	GenBank COI	GenBank 28S	BOLD ID
IU-2008-13009	Norfolk Ridge	*sternomaculata*	holotype	EU243561	EU243662	EUMU225–7
IU-2008-13010	Norfolk ridge	*annulosa*	holotype	EU243515	EU243646	EUMU179–7
IU-2008-13627	Norfolk ridge, Brachiopode	*sternomaculata*		EU243484	EU243635	EUMU148–7
IU-2008-13628	Norfolk ridge, Kaimon Maru	*annulosa*		EU243507	EU243644	EUMU171–7
IU-2008-13642	Norfolk ridge, Introuvable	*sternomaculata*		EU243481	EU243633	EUMU145–7
IU-2008-13736	Norfolk ridge, Kaimon Maru	*annulosa*		EU243506	EU243643	EUMU170–7
IU-2008-13747	Norfolk ridge, Jumeau est	*annulosa*		EU243469	EU243623	EUMU133–7
IU-2008-13748	Norfolk ridge, Jumeau est	*annulosa*		EU243470		EUMU134–7
IU-2008-13749	Norfolk ridge, Eponge	*annulosa*		EU243467	EU243621	EUMU131–7
IU-2008-13750	Norfolk ridge, Eponge	*annulosa*		EU243468	EU243622	EUMU132–7
IU-2008-13751	Norfolk ridge, Crypthelia	*annulosa*		EU243460		EUMU124–7
IU-2008-13752	Norfolk ridge, Crypthelia	*annulosa*		EU243461	EU243618	EUMU125–7
IU-2008-13753	Norfolk ridge, Crypthelia	*annulosa*		EU243462		EUMU126–7
IU-2008-13754	Norfolk ridge, Crypthelia	*annulosa*		EU243463	EU243619	EUMU127–7
IU-2008-13755	Norfolk ridge, Crypthelia	*annulosa*		EU243464		EUMU128–7
IU-2008-13756	Norfolk ridge, Crypthelia	*annulosa*		EU243465		EUMU129–7
IU-2008-13757	Norfolk ridge, Crypthelia	*annulosa*		EU243466	EU243620	EUMU130–7
IU-2008-13758	Norfolk ridge, Introuvable	*annulosa*		EU243474	EU243626	EUMU138–7
IU-2008-13759	Norfolk ridge, Introuvable	*annulosa*		EU243473	EU243625	EUMU137–7
IU-2008-13760	Norfolk ridge, Stylaster	*annulosa*		EU243471	EU243624	EUMU135–7
IU-2008-13761	Norfolk ridge, Stylaster	*annulosa*		EU243472		EUMU136–7
IU-2008-13762	Norfolk ridge, Jumeau ouest	*annulosa*		EU243489		EUMU153–7
IU-2008-13763	Norfolk ridge, Jumeau ouest	*annulosa*		EU243490	EU243637	EUMU154–7
IU-2008-13764	Norfolk ridge, Jumeau ouest	*annulosa*		EU243491		EUMU155–7
IU-2008-13765	Norfolk ridge, Jumeau ouest	*annulosa*		EU243492		EUMU156–7
IU-2008-13766	Norfolk ridge, Jumeau ouest	*annulosa*		EU243493	EU243638	EUMU157–7
IU-2008-13767	Norfolk ridge, Jumeau ouest	*annulosa*		EU243494	EU243639	EUMU158–7
IU-2008-13768	Norfolk ridge, Jumeau est	*sternomaculata*		EU243456	EU243615	EUMU120–7
IU-2008-13769	Norfolk ridge, Jumeau est	*sternomaculata*		EU243457		EUMU121–7
IU-2008-13770	Norfolk ridge, Jumeau est	*sternomaculata*		EU243458	EU243616	EUMU122–7
IU-2008-13771	Norfolk ridge, Jumeau est	*sternomaculata*		EU243459	EU243617	EUMU123–7
IU-2008-13772	Norfolk ridge, Brachiopode	*annulosa*		EU243435		EUMU099–7
IU-2008-13773	Norfolk ridge, Brachiopode	*annulosa*		EU243436		EUMU100–7
IU-2008-13775	Norfolk ridge, Antigonia	*annulosa*		EU243443		EUMU107–7
IU-2008-13776	Norfolk ridge, Antigonia	*annulosa*		EU243444		EUMU108–7
IU-2008-13777	Norfolk ridge, Antigonia	*annulosa*		EU243445		EUMU109–7
IU-2008-13778	Norfolk ridge, Crypthelia	*annulosa*		EU243447		EUMU111–7
IU-2008-13779	Norfolk ridge, Munida	*annulosa*		EU243448		EUMU112–7
IU-2008-13780	Norfolk ridge, Munida	*annulosa*		EU243449	EU243612	EUMU113–7
IU-2008-13781	Norfolk ridge, Munida	*sternomaculata*		EU243450		EUMU114–7
IU-2008-13782	Island of Pines	*annulosa*		EU243451	EU243614	EUMU115–7
IU-2008-13785	Norfolk ridge, Jumeau est	*spinosa*		EU243533	EU243655	EUMU197–7
IU-2008-13786	Norfolk ridge, Jumeau est	*spinosa*		EU243534	EU243656	EUMU198–7
IU-2008-13787	Norfolk ridge, Jumeau est	*spinosa*		EU243535		EUMU199–7
IU-2008-13788	Norfolk ridge, Jumeau est	*spinosa*		EU243536	EU243657	EUMU200–7
IU-2008-13789	Norfolk ridge, Jumeau est	*spinosa*		EU243537		EUMU201–7
IU-2008-13790	Norfolk ridge, Jumeau est	*spinosa*		EU243538		EUMU202–7
IU-2008-13791	Norfolk ridge, Jumeau est	*spinosa*		EU243539		EUMU203–7
IU-2008-13792	Norfolk ridge, Jumeau est	*spinosa*		EU243540		EUMU204–7
IU-2008-13793	Norfolk ridge, Jumeau est	*spinosa*		EU243541		EUMU205–7
IU-2008-13794	Norfolk ridge, Jumeau est	*spinosa*		EU243542		EUMU206–7
IU-2008-13795	Solomon Islands	*laevimana*		EU243508		EUMU172–7
IU-2008-13796	Solomon Islands	*laevimana*		EU243509		EUMU173–7
IU-2008-13797	Guadeloupe	*picta*		EU243556	EU243661	EUMU220–7
IU-2008-13798	Guadeloupe	*picta*		EU243557		EUMU221–7
IU-2008-13799	Guadeloupe	*picta*		EU243558		EUMU222–7
IU-2008-13801	Norfolk ridge, Jumeau est	*sternomaculata*		EU243455		EUMU119–7
IU-2008-13803	Norfolk ridge, Jumeau est	*sternomaculata*		EU243452		EUMU116–7
IU-2008-13804	Norfolk ridge, Jumeau est	*sternomaculata*		EU243453		EUMU117–7

(*continued next page*)

Table 1. (*continued*)

MNHN ID	Geographic area	Morphological ID	Status	GenBank COI	GenBank 28S	BOLD ID
IU-2008-13805	Norfolk ridge, Jumeau est	*sternomaculata*		EU243454		EUMU118–7
IU-2008-13806	Norfolk ridge, Stylaster	*sternomaculata*		EU243400		EUMU064–7
IU-2008-13876	Norfolk ridge, Aztèque	*annulosa*		EU243383		EUMU047–7
IU-2011-5396	Norfolk ridge, Jumeau Ouest	*annulosa*		EU243365		EUMU029–7
IU-2011-5397	Norfolk ridge, Eponge	*annulosa*		EU243370		EUMU034–7
IU-2011-5398	Norfolk ridge, Eponge	*annulosa*		EU243371		EUMU035–7
IU-2011-5399	Norfolk ridge, Eponge	*annulosa*		EU243372		EUMU036–7
IU-2011-5400	Norfolk ridge, Eponge	*annulosa*		EU243412	EU243597	EUMU076–7
IU-2011-5401	Norfolk ridge, Eponge	*annulosa*		EU243413	EU243598	EUMU077–7
IU-2011-5402	Norfolk ridge, Eponge	*annulosa*		EU243414	EU243599	EUMU078–7
IU-2011-5403	Norfolk ridge, Eponge	*annulosa*		EU243415		EUMU079–7
IU-2011-5404	Norfolk ridge, Eponge	*annulosa*		EU243416		EUMU080–7
IU-2011-5405	Norfolk ridge, Eponge	*annulosa*		EU243417		EUMU081–7
IU-2011-5406	Norfolk ridge, Jumeau Ouest	*annulosa*		EU243366	EU243585	EUMU030–7
IU-2011-5407	Norfolk ridge, Jumeau Ouest	*annulosa*		EU243367	EU243586	EUMU031–7
IU-2011-5408	Norfolk ridge, Jumeau Ouest	*annulosa*		EU243368		EUMU032–7
IU-2011-5409	Norfolk ridge, Eponge	*sternomaculata*		EU243475	EU243627	EUMU139–7
IU-2011-5410	Norfolk ridge, Eponge	*sternomaculata*		EU243476	EU243628	EUMU140–7
IU-2011-5411	Norfolk ridge, Eponge	*sternomaculata*		EU243477	EU243629	EUMU141–7
IU-2011-5412	Norfolk ridge, Stylaster	*sternomaculata*		EU243478	EU243630	EUMU142–7
IU-2011-5413	Norfolk ridge, Stylaster	*sternomaculata*		EU243480	EU243632	EUMU144–7
IU-2011-5414	Norfolk ridge, Stylaster	*sternomaculata*		EU243479	EU243631	EUMU143–7
IU-2011-5415	Norfolk ridge, Brachiopode	*sternomaculata*		EU243482	EU243634	EUMU146–7
IU-2011-5416	Norfolk ridge, Brachiopode	*sternomaculata*		EU243483		EUMU147–7
IU-2011-5417	Norfolk ridge, Brachiopode	*sternomaculata*		EU243485		EUMU149–7
IU-2011-5418	Norfolk ridge, Brachiopode	*sternomaculata*		EU243486		EUMU150–7
IU-2011-5419	Norfolk ridge, Brachiopode	*sternomaculata*		EU243487		EUMU151–7
IU-2011-5420	Norfolk ridge, Brachiopode	*sternomaculata*		EU243488	EU243636	EUMU152–7
IU-2011-5421	Norfolk ridge	*sternomaculata*	paratype		EU243651	
IU-2011-5422	Norfolk ridge	*sternomaculata*	paratype		EU243652	
IU-2011-5423	Tuamotu	*keijii*		EU243337		EUMU001–7
IU-2011-5424	New-Caledonia	*keijii*		EU243338		EUMU002–7
IU-2011-5425	New-Caledonia	*keijii*		EU243339		EUMU003–7
IU-2011-5426	Wallis	*keijii*		EU243340		EUMU004–7
IU-2011-5427	New-Caledonia	*capillata*		EU243341		EUMU005–7
IU-2011-5428	New-Caledonia	*capillata*		EU243342		EUMU006–7
IU-2011-5429	Indonesia, Kai Island	*capillata*		EU243343		EUMU007–7
IU-2011-5430	Indonesia, Tanimbar Island	*capillata*		EU243344		EUMU008–7
IU-2011-5431	New-Caledonia, Surprise	*parva*		EU243345		EUMU009–7
IU-2011-5432	New-Caledonia, Surprise	*parva*		EU243346		EUMU010–7
IU-2011-5433	Norfolk Ridge, Jumeau Est	*karubar*		EU243347	EU243574	EUMU011–7
IU-2011-5434	Norfolk Ridge, Jumeau Est	*karubar*		EU243348		EUMU012–7
IU-2011-5435	Norfolk Ridge, Jumeau Est	*karubar*		EU243349	EU243575	EUMU013–7
IU-2011-5436	Indonesia, Tanimbar and Kai Islands	*smithii*		EU243350		EUMU014–7
IU-2011-5437	Indonesia, Tanimbar and Kai Islands	*smithii*		EU243351		EUMU015–7
IU-2011-5438		*treguieri*		EU243352	EU243576	EUMU016–7
IU-2011-5439	Norfolk ridge, Introuvable	*annulosa*		EU243353	EU243577	EUMU017–7
IU-2011-5440	Norfolk ridge, Introuvable	*annulosa*		EU243354	EU243578	EUMU018–7
IU-2011-5441	Norfolk ridge, Introuvable	*sternomaculata*		EU243355		EUMU019–7
IU-2011-5442	Norfolk ridge, Introuvable	*sternomaculata*		EU243356	EU243579	EUMU020–7
IU-2011-5443	Norfolk ridge, Introuvable	*annulosa*		EU243357	EU243580	EUMU021–7
IU-2011-5444	Polynesia, Raivavae	*treguieri*	paratype	EU243358	EU243581	EUMU022–7
IU-2011-5445	Tuamotu, Mururoa	*treguieri*	paratype	EU243359	EU243582	EUMU023–7
IU-2011-5446	Norfolk ridge, Introuvable	*annulosa*		EU243360	EU243583	EUMU024–7
IU-2011-5447	Norfolk ridge, Stylaster	*annulosa*		EU243361	EU243584	EUMU025–7
IU-2011-5448	Norfolk ridge, Stylaster	*annulosa*		EU243362		EUMU026–7
IU-2011-5449	Norfolk ridge, Stylaster	*annulosa*		EU243363		EUMU027–7
IU-2011-5450	Norfolk ridge, Stylaster	*annulosa*		EU243364		EUMU028–7
IU-2011-5451	Norfolk ridge, Eponge	*annulosa*		EU243369		EUMU033–7

(*continued next page*)

Table 1. (*continued*)

MNHN ID	Geographic area	Morphological ID	Status	GenBank COI	GenBank 28S	BOLD ID
IU-2011-5452	Norfolk ridge, Introuvable	*annulosa*		EU243373		EUMU037–7
IU-2011-5453	Norfolk ridge, Introuvable	*annulosa*		EU243374		EUMU038–7
IU-2011-5454	Norfolk ridge, Introuvable	*annulosa*		EU243375		EUMU039–7
IU-2011-5455	Norfolk ridge, Introuvable	*annulosa*		EU243376		EUMU040–7
IU-2011-5456	Norfolk ridge, Jumeau Est	*annulosa*		EU243377		EUMU041–7
IU-2011-5457	Norfolk ridge, Jumeau Est	*annulosa*		EU243378		EUMU042–7
IU-2011-5458	Norfolk ridge, Jumeau Est	*annulosa*		EU243379		EUMU043–7
IU-2011-5459	Norfolk ridge, Jumeau Est	*annulosa*		EU243380		EUMU044–7
IU-2011-5460	Norfolk ridge, Mont n°2	*annulosa*		EU243381		EUMU045–7
IU-2011-5461	Norfolk ridge, Mont n°2	*annulosa*		EU243382		EUMU046–7
IU-2011-5462	Norfolk ridge, Mont n°2	*sternomaculata*		EU243384	EU243587	EUMU048–7
IU-2011-5463	Norfolk ridge, Mont n°1	*sternomaculata*		EU243385	EU243588	EUMU049–7
IU-2011-5464	Norfolk ridge, Mont n°1	*sternomaculata*		EU243386	EU243589	EUMU050–7
IU-2011-5465	Norfolk ridge, Mont n°1	*sternomaculata*		EU243387		EUMU051–7
IU-2011-5466	Norfolk ridge, Mont n°1	*sternomaculata*		EU243388	EU243590	EUMU052–7
IU-2011-5467	Norfolk ridge, Stylaster	*sternomaculata*		EU243389		EUMU053–7
IU-2011-5468	Norfolk ridge, Stylaster	*sternomaculata*		EU243390		EUMU054–7
IU-2011-5469	Norfolk ridge, Stylaster	*sternomaculata*		EU243391		EUMU055–7
IU-2011-5470	Norfolk ridge, Stylaster	*sternomaculata*		EU243392		EUMU056–7
IU-2011-5471	Norfolk ridge, Eponge	*sternomaculata*		EU243393		EUMU057–7
IU-2011-5472	Norfolk ridge, Eponge	*sternomaculata*		EU243394		EUMU058–7
IU-2011-5473	Norfolk ridge, Eponge	*sternomaculata*		EU243395		EUMU059–7
IU-2011-5474	Norfolk ridge, Introuvable	*sternomaculata*		EU243396	EU243591	EUMU060–7
IU-2011-5475	Norfolk ridge, Introuvable	*sternomaculata*		EU243397		EUMU061–7
IU-2011-5476	Norfolk ridge, Introuvable	*sternomaculata*		EU243398		EUMU062–7
IU-2011-5477	Norfolk ridge, Introuvable	*sternomaculata*		EU243399		EUMU063–7
IU-2011-5478	Norfolk ridge, Stylaster	*annulosa*		EU243401		EUMU065–7
IU-2011-5479	Norfolk ridge, Stylaster	*annulosa*		EU243402	EU243592	EUMU066–7
IU-2011-5480	Norfolk ridge, Stylaster	*annulosa*		EU243403		EUMU067–7
IU-2011-5481	Norfolk ridge, Stylaster	*annulosa*		EU243404	EU243593	EUMU068–7
IU-2011-5482	Norfolk ridge, Stylaster	*annulosa*		EU243405		EUMU069–7
IU-2011-5483	Norfolk ridge, Stylaster	*annulosa*		EU243406		EUMU070–7
IU-2011-5484	Norfolk ridge, Introuvable	*annulosa*		EU243407	EU243594	EUMU071–7
IU-2011-5485	Norfolk ridge, Introuvable	*annulosa*		EU243408		EUMU072–7
IU-2011-5486	Norfolk ridge, Introuvable	*annulosa*		EU243409		EUMU073–7
IU-2011-5487	Norfolk ridge, Introuvable	*annulosa*		EU243410	EU243595	EUMU074–7
IU-2011-5488	Norfolk ridge, Introuvable	*annulosa*		EU243411	EU243596	EUMU075–7
IU-2011-5489	Norfolk ridge, Jumeau Est	*annulosa*		EU243418	EU243600	EUMU082–7
IU-2011-5490	Norfolk ridge, Jumeau Est	*annulosa*		EU243419		EUMU083–7
IU-2011-5491	Norfolk ridge, Jumeau Est	*annulosa*		EU243420	EU243601	EUMU084–7
IU-2011-5492	Norfolk ridge, Jumeau Est	*annulosa*		EU243421		EUMU085–7
IU-2011-5493	Norfolk ridge, Jumeau Est	*annulosa*		EU243422	EU243602	EUMU086–7
IU-2011-5494	Norfolk ridge, Stylaster	*sternomaculata*		EU243423	EU243603	EUMU087–7
IU-2011-5495	Norfolk ridge, Stylaster	*sternomaculata*		EU243424		EUMU088–7
IU-2011-5496	Norfolk ridge, Stylaster	*sternomaculata*		EU243425	EU243604	EUMU089–7
IU-2011-5497	Norfolk ridge, Stylaster	*sternomaculata*		EU243426		EUMU090–7
IU-2011-5498	Norfolk ridge, Stylaster	*sternomaculata*		EU243427	EU243605	EUMU091–7
IU-2011-5499	Norfolk ridge, Stylaster	*sternomaculata*		EU243428		EUMU092–7
IU-2011-5500	Norfolk ridge, Eponge	*sternomaculata*		EU243429	EU243606	EUMU093–7
IU-2011-5501	Norfolk ridge, Eponge	*sternomaculata*		EU243430		EUMU094–7
IU-2011-5502	Norfolk ridge, Eponge	*sternomaculata*		EU243431		EUMU095–7
IU-2011-5503	Norfolk ridge, Eponge	*sternomaculata*		EU243432	EU243607	EUMU096–7
IU-2011-5504	Norfolk ridge, Eponge	*sternomaculata*		EU243433		EUMU097–7
IU-2011-5505	Norfolk ridge, Eponge	*sternomaculata*		EU243434	EU243608	EUMU098–7
IU-2011-5506	Norfolk ridge, Brachiopode	*annulosa*		EU243437	EU243609	EUMU101–7
IU-2011-5507	Norfolk ridge, Brachiopode	*annulosa*		EU243438	EU243610	EUMU102–7
IU-2011-5508	Norfolk ridge, Kaimon Maru	*annulosa*		EU243439		EUMU103–7
IU-2011-5509	Norfolk ridge, Kaimon Maru	*annulosa*		EU243440		EUMU104–7
IU-2011-5510	Norfolk ridge, Kaimon Maru	*annulosa*		EU243441		EUMU105–7

(*continued next page*)

Table 1. (*continued*)

MNHN ID	Geographic area	Morphological ID	Status	GenBank COI	GenBank 28S	BOLD ID
IU-2011-5511	Norfolk ridge, Kaimon Maru	*annulosa*		EU243442	EU243611	EUMU106–7
IU-2011-5512	Norfolk ridge, Crypthelia	*annulosa*		EU243446		EUMU110–7
IU-2011-5513	Norfolk ridge, Jumeau ouest	*annulosa*		EU243495		EUMU159–7
IU-2011-5514	Norfolk ridge, Jumeau ouest	*annulosa*		EU243496		EUMU160–7
IU-2011-5515	Solomon Islands	*laevimana*		EU243497		EUMU161–7
IU-2011-5516	Madagascar	*similior*	holotype	EU243498		EUMU162–7
IU-2011-5517	Indonesia, Tanimbar and Kai Islands	*treguieri*		EU243499		EUMU163–7
IU-2011-5518	Norfolk ridge, Jumeau est	*spinosa*	paratype	EU243500	EU243640	EUMU164–7
IU-2011-5519	Norfolk ridge, Jumeau est	*spinosa*	paratype	EU243501	EU243641	EUMU165–7
IU-2011-5520	Loyalty ridge	*minor*		EU243502		EUMU166–7
IU-2011-5521	Norfolk ridge, Introuvable	*sternomaculata*		EU243503		EUMU167–7
IU-2011-5522	Norfolk ridge, Introuvable	*sternomaculata*		EU243504	EU243642	EUMU168–7
IU-2011-5523	Norfolk ridge, Introuvable	*sternomaculata*		EU243505		EUMU169–7
IU-2011-5524	Solomon Islands	*laevimana*		EU243510		EUMU174–7
IU-2011-5525	Polynesia, Tubuai	*treguieri*	paratype	EU243511		EUMU175–7
IU-2011-5526	Polynesia, Tubuai	*treguieri*	paratype	EU243512		EUMU176–7
IU-2011-5527	Norfolk ridge, Jumeau est	*spinosa*	holotype	EU243513	EU243645	EUMU177–7
IU-2011-5528	New-Caledonia	*keijii*	paratype	EU243514		EUMU178–7
IU-2011-5529	Tuamotu	*treguieri*	paratype	EU243516	EU243647	EUMU180–7
IU-2011-5530	Tuamotu	*treguieri*	paratype	EU243517	EU243648	EUMU181–7
IU-2011-5531	New-Caledonia	*parva*	holotype	EU243518		EUMU182–7
IU-2011-5532	Norfolk ridge, Stylaster	*annulosa*		EU243519	EU243649	EUMU183–7
IU-2011-5533	Norfolk ridge, Stylaster	*annulosa*		EU243520	EU243650	EUMU184–7
IU-2011-5534	New-Caledonia, Surprise	*parva*		EU243521		EUMU185–7
IU-2011-5535	New-Caledonia, Surprise	*parva*		EU243522	EU243653	EUMU186–7
IU-2011-5536	New-Caledonia, Surprise	*parva*		EU243523		EUMU187–7
IU-2011-5537	New-Caledonia, Surprise	*parva*		EU243524		EUMU188–7
IU-2011-5538	New-Caledonia, Surprise	*parva*		EU243525	EU243654	EUMU189–7
IU-2011-5539	New-Caledonia, Surprise	*parva*		EU243526		EUMU190–7
IU-2011-5540	New-Caledonia, Surprise	*parva*		EU243527		EUMU191–7
IU-2011-5541	Indonesia, Kai island	*karubar*	paratype	EU243528		EUMU192–7
IU-2011-5542	Indonesia, Kai island	*karubar*	paratype	EU243529		EUMU193–7
IU-2011-5543	Indonesia, Kai island	*karubar*	paratype	EU243530		EUMU194–7
IU-2011-5544	Indonesia, Kai island	*karubar*	paratype	EU243531		EUMU195–7
IU-2011-5545	Indonesia, Kai island	*karubar*	paratype	EU243532		EUMU196–7
IU-2011-5546	New-Caledonia	*marginata*	holotype	EU243543		EUMU207–7
IU-2011-5547	Madagascar	*bispinata*	paratype	EU243544		EUMU208–7
IU-2011-5548	Madagascar	*bispinata*	paratype	EU243545		EUMU209–7
IU-2011-5549	Madagascar	*multilineata*	paratype	EU243546		EUMU210–7
IU-2011-5550	Madagascar	*minor*		EU243547		EUMU211–7
IU-2011-5551	Norfolk ridge	*minor*	paratype	EU243548		EUMU212–7
IU-2011-5552	Norfolk ridge	*minor*	paratype	EU243549		EUMU213–7
IU-2011-5553	Loyalty ridge	*minor*	holotype	EU243550		EUMU214–7
IU-2011-5554	Loyalty ridge	*minor*	paratype	EU243551		EUMU215–7
IU-2011-5555	Loyalty ridge	*minor*		EU243552		EUMU216–7
IU-2011-5556	Loyalty ridge	*minor*		EU243553		EUMU217–7
IU-2011-5557	Loyalty ridge	*minor*		EU243554		EUMU218–7
IU-2011-5558	Philippines	*funambulus*			EU243658	
IU-2011-5559	Indonesia, Kai Island	*karubar*	holotype	EU243555	EU243659	EUMU219–7
IU-2011-5560		*treguieri*			EU243660	
IU-2011-5561	Namibia	*squamifera*	paratype	EU243559		EUMU223–7
IU-2011-5562	Namibia	*squamifera*	paratype	EU243560		EUMU224–7
IU-2011-5563	Polynesia, Tuamotu	*treguieri*	holotype	EU243562	EU243663	EUMU226–7

calculated for the COI gene under the K2P model and used to reconstruct a Neighbour-Joining tree, using MEGA 5 (Tamura *et al*. 2011). To accurately reconstruct the phylogenetic relationships within *Eumunida*, a Bayesian Analysis was also conducted using Mr. Bayes (Huelsenbeck *et al*. 2001); it consisted of two independent analyses (six Markov chains, 30 000 000 generations, with a sampling frequency of one tree each 5000 generations). One different model (each with 6 substitution categories, a gamma-distributed rate variation across sites approximated in four discrete categories and a proportion of invariable sites) was applied for each partition (COI, 28S loops and 28S stems). Convergence of each

Table 2. Description of morphological characters

	Characters	States
1	Thoracic spines	Yes = 1, No = 0
2	Posterior region of carapace with complete striae	Yes = 1, No = 0
3	Number of anterolateral spines on each side	One spine = 1, two spines = 0
4	Pad on palm of cheliped	Yes = 1, No = 0
5	Epigastric spines	Yes = 1, No = 0
6	Posterior part of abdominal tergites, after last stria, smooth	Yes = 1, No = 0
7	Depressed area on branchial region of carapace	Yes = 1, No = 0
8	Mesiodorsal row of spines on cheliped palm	Yes = 1, No = 0
9	First anterolateral spine less than half lateral supraorbital	Yes = 1 (less), No = 0 (more)
10	Distal spines on carpus of chelipeds	2 spp. = 1, 3 spp. = 0
11	Distal spine on merus of third maxilliped	Yes = 1, No = 0
12	Male pleopods	Yes = 1, No = 0
13	Six to seven spines on upper margin of propodus walking leg	Yes = 1, No = 0
14	Row of ventral spines on merus of chelipeds	5–8 spp. = 1, 1 sp. = 0
15	Ocular peduncles short, not reaching end of lateral supraorbital spines	Yes = 1, No = 0
16	Lateral surface of 4th pereiopod with spine	Yes = 1, No = 0

Table 3. Character state for each species of *Eumunida*

Species for which molecular data were obtained are shown in bold. See Table 2 for explanation of the morphological characters used

	Characters															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
E. ampliata	0	1	1	1	0	1	0	0	1	0	0	0	0	1	0	1
E. annulosa	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	1
E. australis	1	1	0	1	0	1	0	0	0	0	0	0	0	1	1	1
E. balssi	0	1	0	0	0	1	0	1	1	0	1	0	0	0	0	1
E. bella	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	1
E. bispinata	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1
E. capillata	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0
E. chani	1	1	0	0	0	1	0	0	1	1	0	0	0	1	0	1
E. debilistriata	0	0	1	0	0	0	0	1	0	0	0	0	1	1	0	1
E. depressa	1	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0
E. dofleini	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
E. funambulus	1	1	1	1	1	0	0	1	0	0	1	0	1	1	0	1
E. gordonae	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0
E. karubar	0	1	0	0	0	0	0	1	1	0	0	1	0	1	0	1
E. keijii	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1
E. laevimana	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0
E. macphersoni	1	1	1	0	0	0	0	1	1	0	0	0	0	0	1	0
E. marginata	0	1	0	1	1	0	0	1	1	0	1	0	0	0	0	0
E. minor	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1
E. multilineata	1	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0
E. pacifica	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1
E. parva	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1
E. picta	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1
E. similior	1	1	1	1	0	1	0	0	0	0	0	0	0	1	0	0
E. smithii	0	1	0	0	0	0	0	1	1	0	0	1	0	1	1	1
E. spinosa	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
E. squamifera	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
E. sternomaculata	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
E. treguieri	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1
Eumunida sp. nov.	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1

analysis was evaluated using Tracer 1.4.1 (Rambaut and Drummond 2007), and analyses were terminated when ESS values were all greater than 200. We also used the AWTY application (a system for graphical exploration of Markov Chain Monte Carlo convergence in Bayesian phylogenetic inference) for each run (two runs for the COI genes and two for the 28S gene): the cumulative split frequencies were stable after the burn-in phase, the split frequencies in run pairs

('compare' analysis) were strongly correlated and the between-run distance was included in the range of the within-run distances for more than half of the generations ('var' analysis). A consensus tree was then calculated after omitting the first 25% trees as burn-in. For both genes, we used *Munida acantha* (Macpherson, 1994) as an outgroup to artificially root the tree (GenBank accession numbers: AY800033 for COI gene and EU249347 for 28S rDNA gene).

Results

Mitochondrial dataset

We obtained 226 COI sequences of 658 bp in length with 219 polymorphic sites corresponding mostly to the first (47) and third (164) codon position. This dataset is available in the BOLD project 'Eumunida barcodes and taxonomy' under the accession numbers EUMU001–07 to EUMU226–07. The maximum K2P distance between pairs of COI sequences of the genus *Eumunida* is 0.158, with a minimum of 0 and a mean of 0.079 (Fig. 1*A*). The histogram representing all the distances between types and non-type specimens defines two groups (Fig. 1*A*): the first, with an upper boundary of 0.033, includes all the distances between two type specimens of one species, but also distances between the holotype of *E. parva* (de Saint Laurent & Macpherson, 1990) and the type specimens (one holotype and five paratypes) of *E. karubar* (de Saint Laurent & Poupin, 1996); the second, characterised by a lower boundary of 0.043, includes only interspecific comparisons between types. Neighbour-Joining and Bayesian phylogenetic trees were highly congruent (only the Bayesian tree is shown in Fig. 2*A*) and revealed 16 terminal genetic units: genetic distances within each cluster are less than 0.033, and COI sequences placed in different genetic units are separated by genetic distances greater than 0.043. Among these 16 genetic units, 13 include several specimens and all are highly supported (Posterior Probabilities pp. = 1), and 10 contain one or several type specimens.

Nuclear dataset

The 28S rDNA gene was much more difficult to sequence, especially for older museum specimens and, as a consequence, fewer specimens were sequenced compared with the CO1 dataset. We obtained 89 sequences of 867 bp. Two groups of K2P distances are separated by a gap on the genetic distances histogram (Fig. 1*B*). The short-distance group has an upper bound of 0.001 and the long-distance group has a lower bound of 0.018. For each pair of specimens, a genetic distance less than 0.001 for this dataset corresponds to a genetic distance less than

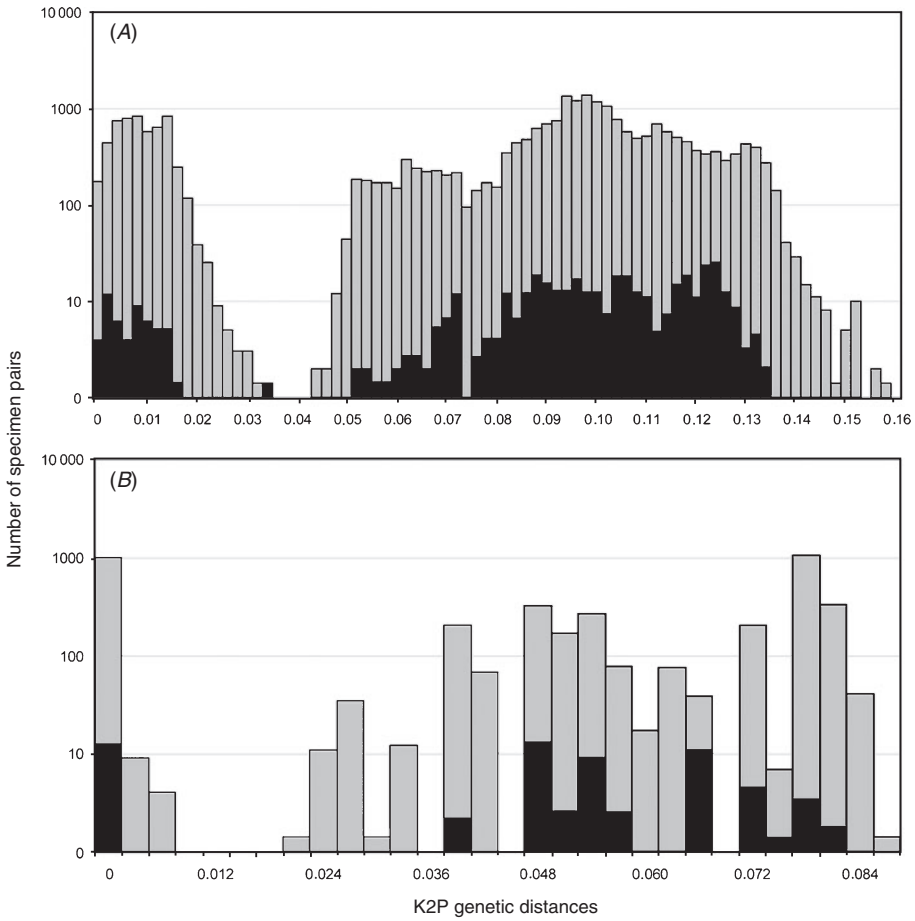

Fig. 1. Histogram of genetic distances for (*A*) the COI gene and (*B*) the 28S gene datasets. Black bars: pairs of type specimens. Grey bars: pairs of non-type specimens.

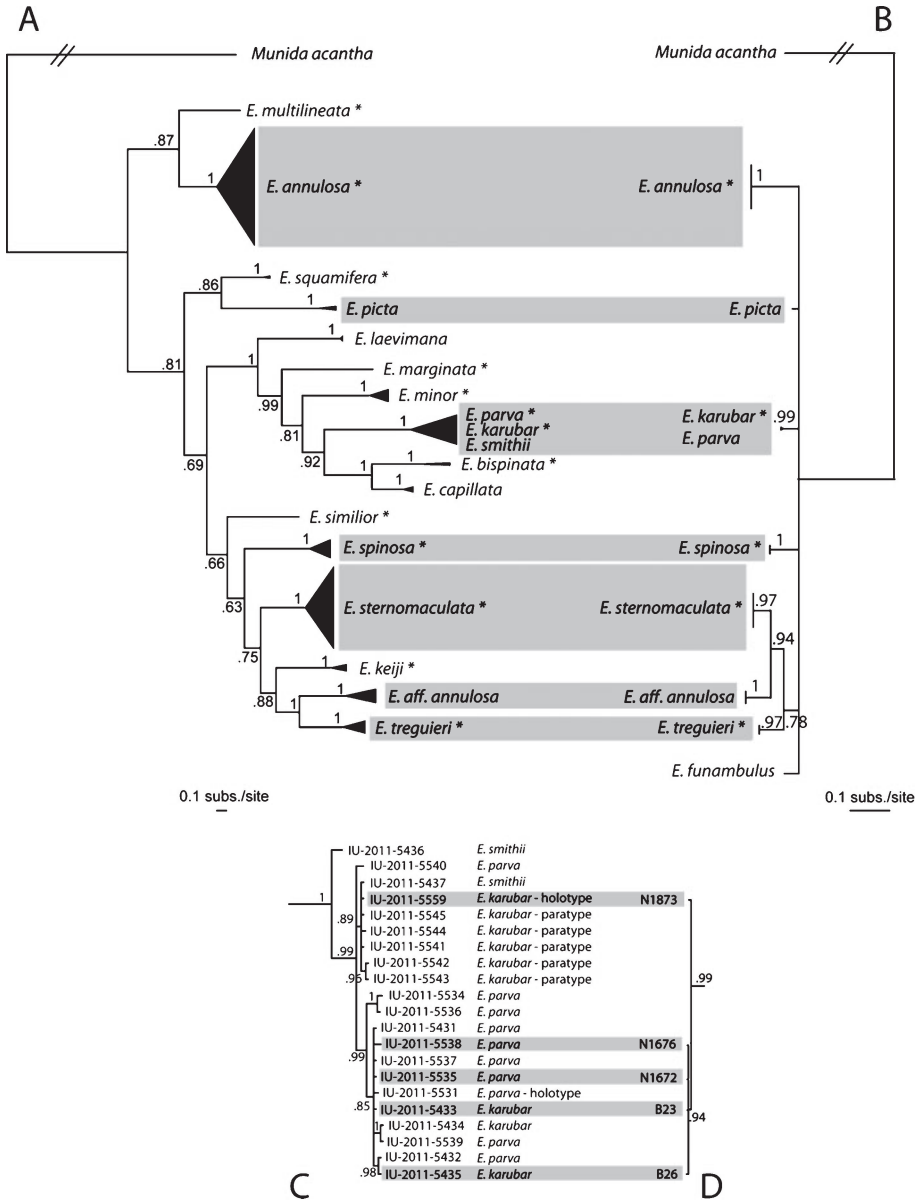

Fig. 2. (*A*) Bayesian tree for COI gene dataset, with posterior probabilities indicated for each node. Clades are collapsed in triangles, with the height representing the number of specimens and the width the length of the branches. An asterisk indicates units that include a type specimen. (*B*) Bayesian tree for the 28S gene dataset. (*C*) Detail of the COI gene tree for the *E. parva*/*E. karubar*/*E. smithii* clade. (*D*) Detail of the 28S gene tree for the *E. parva*/*E. karubar*/*E. smithii* clade.

0.033 with the COI gene. Conversely, when the genetic distance between two 28S rDNA sequences is greater than 0.018, the genetic distance between COI sequences corresponding to the same specimens is greater than 0.043. The intraspecific distances between type specimens fall in the short-distance group whereas interspecific distances between type specimens fall in the long-distance group. The 28S dataset reveals the same monophyletic lineages as the COI dataset: among the 16 lineages defined with the COI gene, seven correspond with clusters identified by the 28S gene (Fig. 2*B*). Furthermore, one additional lineage, not sequenced with the COI gene, is defined with the 28S gene. The deeper nodes of the 28S tree are not as well resolved as the CO1 tree but the terminals are highly supported in all cases.

Genetic units and species names

On the basis of separate analyses of the two molecular datasets, we are able to define 17 genetically distinct units (Fig. 2) that may be considered as species hypotheses. Eleven of these units include at least one sequence of one type specimen (holotype and/or paratypes) for at least one of the two genes, and can be directly linked to a species name. Types were included for *E. annulosa* de

Saint Laurent & Macpherson, 1990, *E. bispinata* Baba, 1990, *E. keiji* de Saint Laurent & Macpherson, 1990, *E. marginata* de Saint Laurent & Macpherson, 1990, *E. minor* de Saint Laurent & Macpherson, 1990, *E. multilineata* de Saint Laurent & Poupin, 1996, *E. similior* Baba, 1990, *E. spinosa* Macpherson, 2006, *E. squamifera* de Saint Laurent & Macpherson, 1990, *E. sternomaculata* de Saint Laurent & Macpherson, 1990 and *E. treguieri* de Saint Laurent & Poupin, 1996. Four other genetic units do not include type specimens but their identification is based on morphological identification keys: *E. capillata* de Saint Laurent & Macpherson, 1990, *E. funambulus* Gordon, 1930, *E. laevimana* Gordon, 1930, and *E. picta* Smith, 1883. The name '*E. annulosa*' is attributed to two clades, one including the holotype. Since the specimens of the genetic unit without the holotype look like those from *E. annulosa* but are not closely related to *E. annulosa* (Fig. 2*A*, *B*), in accordance with the Code of Zoological Nomenclature, we named this genetic group *E.* aff. *annulosa*. Finally, the remaining genetic unit unites specimens morphologically assigned to three different species (*E. karubar* de Saint Laurent & Poupin, 1996; *E. parva* de Saint Laurent & Macpherson, 1990, and *E. smithii* Henderson, 1885). For the COI dataset, the holotype of *E. parva*, five paratypes of *E. karubar* and the holotype of *E. karubar* are included within the same genetic unit (Fig. 2*C*). Genetic distances between sequences of paratypes and/or holotypes falling into this well supported clade are lower than between other paratypes of a single species name placed in a single clade (e.g. the two paratypes of *E. bispinata*).

Discussion

The barcoding gap

In our analysis, the distribution pattern of genetic distances for the two gene fragments used allows us to cluster genetically similar individuals that are separated from each other by relatively large distances. In the bimodal distribution of distances, the lower bound of the first mode – small distances – and the upper bound of the second mode – large distances (Meier *et al*. 2008) – are reliably estimated thanks to the larger number of specimens analysed, allowing the assertion that the observed gap is not an artefact resulting from a sampling bias. We are fully aware, like others (e.g. Meyer and Paulay 2005; Costa *et al*. 2007; Hajibabaei *et al*. 2007; Wiemers and Fiedler 2007; Meier *et al*. 2008), of the importance of the sampling scheme to interpret a gap in the distribution of the pairwise genetic distances, but insist that the originality of our dataset is the inclusion of type specimens. Interestingly, all the genetic distances between the paratypes of a given name fall in the first mode whereas genetic distances among the holotypes (and the paratypes from different names) fall into the second mode (except for the type specimens of *E. karubar* and *E. parva*), suggesting that the gap may be used in a first approach as a species threshold.

Concordance of most genetic units with primary species hypotheses

Inclusion of a closely related outgroup in the analysis shows that each of the 17 defined genetic units has it own evolutionary history. Moreover, the two gene trees obtained with our two unlinked genetic markers are in concordance. This concordance suggests that genetic exchanges among individuals from different clades are unlikely. A previous study has shown that in two of these genetic units, gene flow occurs between populations over the geographic range of each species but not between species (Samadi *et al*. 2006). These 17 genetic units can thus be considered robust species hypotheses.

Among them, 15 units cluster specimens attributed to a unique species and a single name using the morphological identification key. Ten of these 15 species clusters also include type specimens. These 15 clusters are therefore delimited unambiguously, even though inclusion of type specimens in such genetic units is the only way to unambiguously attribute species names to them; but even though five units do not include the type specimen for the name attributed from the key, we can define 15 primary species hypotheses as the best ones given the available data to date. However, our result is not fully congruent with previous species hypotheses, of which four are questioned by the molecular analysis. Indeed, our data suggest (1) the occurrence of a cryptic species (i.e. not yet identified using morphology) that needs a new name because no type specimen can be attributed to the corresponding cluster, and (2) the grouping of three previously admitted species hypotheses into one, and thus the synonymy of three available species names.

A cryptic species under the name E. annulosa

The genetic divergence found between *E. annulosa* and *E.* aff. *annulosa* largely exceeds the average divergence found not only within the other species hypotheses of our dataset, but also within other galatheoid species (Machordom and Macpherson 2004). Since one of the two clades includes the holotype of *E. annulosa*, the other clade (*E.* aff. *annulosa*), not yet detected by morphologists, should indisputably be described under a new name (Fig. 2*A*, *B*). Although this clade is more closely related to *E. treguieri* in the tree, the morphological characters differ only slightly from those of *E. annulosa* or *E. sternomaculata*. These two species are distinguished morphologically by the relative length of the first pair of anterolateral spines (longer in *E. sternomaculata* than in *E. annulosa*), the presence of two (*E. annulosa*) or three (*E. sternomaculata*) distal spines on the carpus of the chelipeds, and the posterior part of the abdominal tergites, after last stria (smoother in *E. annulosa* than in *E. sternomaculata*) (Table 3: Characters 6, 9 and 10). The larger specimens of *E.* aff. *annulosa* display intermediate states for two characters: the relative size of the first anterolateral spine is intermediate between that described for *E. annulosa* and that described for *E. sternomaculata* and a 3rd distal spine is present on the cheliped carpus, but is generally very small. However, these morphological characters, on which this new species may be diagnosed, are difficult to observe on small specimens and thus are useful only for identification of adult specimens. Since the two species are morphologically very close but do not display sister relationships, they are 'cryptic species', and not 'sibling species', as defined by Bickford *et al*. (2007). This result stresses the importance of molecular analyses to detect such 'cryptic species', not only within this genus but also in other crustacean decapods (see

the review by Knowlton 2000 and Bickford *et al*. 2007). Contrary to most studies, which provide (at best) molecular data for name-bearing specimens of new species names (e.g. Shih *et al*. 2010; Ahyong *et al*. 2010), the inclusion of many name-bearing specimens in the analysis points to the necessity of a new name for this 'cryptic species'. For such cryptic taxa, the DNA barcode is obviously a more effective identification tool than a morphological identification key, being informative at all life stages and thus having broader applications (e.g. De Ley *et al*. 2005; Savolainen *et al*. 2005; Vences *et al*. 2005).

Synonymy of E. karubar, E. parva and E. smithii

Our analysis also suggests that three named species hypotheses (*E. karubar*, *E. parva*, *E. smithii*) should actually be merged into a single species hypothesis. When using a morphological identification key, the specimens attributed to each one of these three species names, including the five paratypes, the holotype of *E. karubar* and the holotype of *E. parva*, are scattered among the different subclades without showing any obvious significant pattern (Fig. 2*C*, *D*).

The morphological distinction among *E. parva*, *E. karubar* and *E. smithii* is based on the occurrence (*E. smithii* and *E. karubar*) or absence (*E. parva*) of ventral spines on the merus of the chelipeds and on the presence (*E. smithii* and *E. karubar*) or absence (*E. parva*) of some ventromesial spines on the palm of the chelipeds (Table 3: Characters 8 and 14). The distinction among these species is also based on the length of the ocular peduncles (shorter in *E. smithii* than in *E. karubar* and *E. parva*) (Saint Laurent and Poupin 1996). By combining data from morphology, geography, and independent genetic characters, we suggest that the three names are synonymous (this amounts to considering *E. parva* and *E. karubar* as junior synonyms of *E. smithii*). This interpretation may yet be challenged by the molecular analysis of the holotype of *E. smithii*. Such an analysis could not be conducted for this study because the type specimens for this name were collected during the Challenger Expedition (1874–76), are not housed at the MNHN, and tissue was not available for sequencing. Consequently, we used topotypic specimens collected from the type locality (Kei Islands, Indonesia). According to our interpretation, the morphological differences upon which description of new species hypotheses bearing new species names has been based in the past are the expression of intraspecific variability. This would imply that variability should be used with caution as a diagnostic trait at species level in this genus. The alternative hypothesis would be recent speciation events leading to low genetic divergence.

Therefore, we propose that the genus *Eumunida* contains 28 species (see also Baba *et al*. 2008; Schnabel and Ahyong 2010), including the new cryptic species of *E. annulosa* and considering *E. parva* and *E. karubar* as junior synonyms of *E. smithii*). The diagnosis of *E. smithii* is as follows:

Diagnosis of E. smithii

Carapace with distinct transverse ridges, laterally armed with 6 spines; 2 spines anterior to posterior cervical groove, anterior spine subequal to posterior spine, about half as long as lateral supraocular spine. No spine on gastric region. Third maxilliped merus with median spine and without distal spine on extensor margin. Sternite 3 with paired median spines; Sternite 4 unarmed on each side. Cheliped carpus with 3 terminal spines; palm without ventral pad of densely packed hairs, longer than fingers, relatively massive, covered with short fine setae. Rudimentary pleopods present on abdominal segments 2–5 in males.

Name-bearing specimens integrated into a molecular revision of species hypotheses

One of the main problems when revising species hypotheses and identifying specimens in the context of DNA-barcoding projects is the naming procedure. An appropriate sampling effort within species, a large taxonomic coverage within the genus, and the inclusion of as many type specimens as possible are necessary when confronting morphological species hypotheses to independent characters (DNA polymorphism) and various species delimitation criteria. In the case of the genus *Eumunida*, it allowed us (1) to support most of the morphology-based primary species hypotheses, (2) to bring up new hypotheses, and (3) to point to the necessity of a taxonomic revision. Overall, although we detected two discrepancies between our data and the current state of the taxonomy of *Eumunida*, our results suggest that most morphological traits commonly used in this genus to propose primary species hypotheses stand up when other characters are used. By contrast with most studies, the inclusion of name-bearing specimens in the molecular study allows us to correctly assign names to the supported or reformulated species hypotheses and to unquestionably determine whether new names are needed or whether some names should be considered synonyms of older names. This point is particularly critical when cryptic species are detected, i.e. when morphological keys do not help to attribute names to genetic units. Last, even though several *Eumunida* species are missing in this study and should be barcoded in the future, our study shows that the COI gene fragment is an effective tool to attribute species names to specimens, and *vice versa*, in *Eumunida*, which is the primary purpose of DNA barcoding.

Acknowledgements

We are grateful to Bertrand Richer de Forges, cruise leader of several deep-sea cruises of the Tropical Deep Sea Benthos program on board R/V Alis, that generated the deep-sea samples used in this study. All material has been collected under appropriate collection permits and approved ethics guidelines. This work was supported by the 'Consortium National de Recherche en Génomique', and the 'Service de Systematique Moléculaire' of the Muséum national d'Histoire naturelle (UMS 2700 CNRS-MNHN). It is part of the agreement n°2005/67 between the Genoscope and the Muséum national d'Histoire naturelle on the project 'Macrophylogeny of Life' directed by Guillaume Lecointre. These data fed the MarBol project supported by the Sloan Foundation. We are also pleased to thank Anouk Barberousse, Philippe Bouchet, Mande Holford, Line Le Gall and Simon Tillier for constructive discussions and/or comments on, and English improvements of, the manuscript, and Julien Brisset and Laure Corbari for the curation of the collection.

References

Ahyong, S. T., Chan, T.-Y., and Bouchet, P. (2010). Mighty claws: a new genus and species of lobster from the Philippine deep sea (Crustacea, Decapoda, Nephropidae). *Zoosystema* **32**(3), 525–535. doi:10.5252/z2010n3a11

Baba, K., Macpherson, E., Poore, G. C. B., Ahyong, S. T., Bermudez, A., Cabezas, P., Lin, C. W., Nizinski, M., Rodrigues, C., and Schnabel, K. E. (2008). Catalogue of squat lobsters of the world (Crustacea: Decapoda: Anomura – families Chirostylidae, Galatheidae and Kiwaidae). *Zootaxa* **1905**, 1–220.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., and Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution* **22**(3), 148–155. doi:10.1016/j.tree.2006.11.004

Bouchet, P., Héros, V., Lozouet, P., and Maestrati, P. (2008). A quater-century of deep-sea malacological exploration in the South and West Pacific: Where do we stand? How far to go? *Mémoires du Muséum national d'Histoire naturelle* **196**, 9–40.

Costa, F. O., deWaard, J. R., Boutillier, J., Ratsnasingham, S., Dooh, R. T., Hajibabaei, M., and Hebert, P. D. N. (2007). Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* **64**, 272–295. doi:10.1139/f07-008

Dayrat, B. (2006). A taxonomic revision of *Paradoris* sea slugs (Mollusca, Gastropoda, Nudibranchia, Doridina). *Zoological Journal of the Linnean Society* **147**, 125–238. doi:10.1111/j.1096-3642.2006.00219.x

Dayrat, B., Tillier, A., Lecointre, G., and Tillier, S. (2001). New clades of Euthyneuran gastropods (Mollusca) from 28S rRNA sequences. *Molecular Phylogenetics and Evolution* **19**(2), 225–235. doi:10.1006/mpev.2001.0926

De Ley, P., Tandingam De Ley, I. T., Morris, K., Abebe, E., Mundo-Ocampo, M., Yoder, M., Heras, J., Waumann, D., Rocha-Olivares, A., Jay Burr, A. H., Baldwin, J. G., and Thomas, W. K. (2005). An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. *Philosophical Transactions of the Royal Society B* **360**, 1945–1958. doi:10.1098/rstb.2005.1726

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.

Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., and Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics* **23**, 167–172. doi:10.1016/j.tig.2007.02.001

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95–98.

Hebert, P. D. N., and Gregory, T. R. (2005). The promise of DNA Barcoding for taxonomy. *Systematic Biology* **54**, 852–859. doi:10.1080/10635150500354886

Huelsenbeck, J. P., Ronquist, F., and Hall, B. (2001). MrBayes: Bayesian inference of phylogeny. *Bioinformatics (Oxford, England)* **17**, 754–755. doi:10.1093/bioinformatics/17.8.754

Jovelin, R., and Justine, J.-L. (2001). Phylogenetic relationships within the Polyopisthocotylean monogeneans (Plathyhelminthes) inferred from partial 28S rDNA sequences. *International Journal for Parasitology* **31**(4), 393–401. doi:10.1016/S0020-7519(01)00114-X

Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J., and McInerney, J. O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that *ad hoc* assumptions for choice of matrix are not justified. *BMC Evolutionary Biology* **6**, 29. doi:10.1186/1471-2148-6-29

Knowlton, N. (2000). Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* **420**, 73–90. doi:10.1023/A:1003933603879

Machordom, A., and Macpherson, E. (2004). Rapid radiation and cryptic speciation in galatheid crabs of the genus *Munida* and related genera in the South West Pacific: molecular and morphological evidence. *Molecular Phylogenetics and Evolution* **33**, 259–279. doi:10.1016/j.ympev.2004.06.001

Meier, R., Zhang, G., and Ali, F. (2008). The use of mean instead of smallest interspecific distances exaggerates the size of the "Barcoding Gap" and leads to misidentification. *Systematic Biology* **57**(5), 809–813. doi:10.1080/10635150802406343

Meyer, C. P., and Paulay, G. (2005). DNA Barcoding: error rates based on comprehensive sampling. *PLoS Biology* **3**(12), e422. doi:10.1371/journal.pbio.0030422

Padial, J. M., and De La Riva, I. (2007). Integrative taxonomists should use and produce DNA barcodes. *Zootaxa* **1586**, 67–68.

Rambaut, A., and Drummond, A. J. (2007). Tracer v1.4. Available from http://beast.bio.ed.ac.uk/Tracer

Saint Laurent, M. d., and Poupin, J. (1996). Crustacea, Anomura: Les espèces indo-ouest pacifiques du genre *Eumunida* Smith, 1880 (Chirostylidae). Description de six espèces nouvelles. In 'Résultats des Campagnes MUSORSTOM, volume 15. Vol. 168'. (Ed. A Crosnier.) pp. 337–385. (Mémoires du Museum national d'Histoire naturelle: Paris.)

Samadi, S., Bottan, L., McPherson, E., Richer De Forges, B., and Boisselier, M. C. (2006). Seamount endemism questioned by the geographic distribution and population genetic structure of marine invertebrates. *Marine Biology* **149**, 1463–1475. doi:10.1007/s00227-006-0306-4

Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K., and Lane, R. (2005). Towards writing the encyclopedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society B* **360**, 1805–1811. doi:10.1098/rstb.2005.1730

Schindel, D. E., and Miller, S. E. (2005). DNA barcoding a useful tool for taxonomists. *Nature* **435**, 17. doi:10.1038/435017b

Schnabel, K. E., and Ahyong, S. T. (2010). A new classification of the Chirostyloidea (Crustacea: Decapoda: Anomura). *Zootaxa* **2687**, 56–64.

Shih, H.-T., Naruse, T., and Ng, P. K. L. (2010). *Uca jocelynae* sp. nov. a new species of fiddler crab (Crustacea: Brachyura: Ocypodidae) from the Western Pacific. *Zootaxa* **2337**, 47–62.

Smith, A. M., Wood, D. M., Janzen, D. H., Hallwachs, W., and Hebert, P. D. N. (2007). DNA barcodes affirm that 16 species of apparently generalist tropical parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4967–4972. doi:10.1073/pnas.0700050104

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Molecular Biology and Evolution* **28**, 2731–2739. doi:10.1093/molbev/msr121

Vences, M., Thomas, M., Bonett, R. M., and Vieites, D. R. (2005). Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B* **360**(1462), 1859–1868. doi:10.1098/rstb.2005.1717

Wiemers, M., and Fiedler, K. (2007). Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology* **4**, 8. doi:10.1186/1742-9994-4-8